**STIC-Biotech/ChemLib**

*166919*

| | |
|---|---|
| **From:** | Riggins, Patrick S. |
| **Sent:** | Tuesday, September 27, 2005 8:58 AM |
| **To:** | STIC-Biotech/ChemLib |
| **Subject:** | 10714212 |
| **Contacts:** | STIC |

Please search SEQ ID NO: 21 and SEQ ID NO: 22 against the commercial protein databases.

Thank you
Patrick Riggins
Examiner
Art Unit 1633
REM 2D60
(571) 272-6102

********************
Searcher:_____
Searcher Phone:_____
Date Searcher Picked up:_____
Date completed:_____
Searcher Prep Time:_____
Online Time:_____

********************
Type of Search
NA#_____ AA#:_____
S/L:_____ Oligomer:_____
Encode/Transl:_____
Structure #:_____Text:____
Inventor:_____ Litigation:___

*******************
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM:_____
WWW/Internet:_____
Other (Specify):_____

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 29, 2005, 19:30:39 ; Search time 118 Seconds
(without alignments)
62.275 Million cell updates/sec

Title: US-10-714-212-21
Perfect score: 91
Sequence: 1 ESEKTASGIVLPDSAKEK 19
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 91 | 100.0 | 19 | 4 | AAU04230 | Aau04230 Chaperoni |
| 2 | 88 | 96.7 | 94 | 5 | ABB49470 | Abb49470 Listeria |
| 3 | 85 | 93.4 | 94 | 5 | AAU04229 | Aau04229 Chaperoni |
| 4 | 85 | 93.4 | 19 | 4 | AAU04231 | Aau04231 Chaperoni |
| 5 | 85 | 93.4 | 40 | 2 | AAR20197 | Aar20197 Heat shoc |
| 6 | 81 | 89.0 | 94 | 2 | ADS21830 | Ads21830 Bacterial |
| 7 | 80 | 87.9 | 19 | 4 | AAU04238 | Aau04238 Chaperoni |
| 8 | 80 | 87.9 | 19 | 4 | AAU04239 | Aau04239 Chaperoni |
| 9 | 80 | 87.9 | 94 | 4 | AAR67388 | Aar67388 GroES-lik |
| 10 | 78 | 85.7 | 95 | 8 | ADS28875 | Ads28875 Bacterial |
| 11 | 78 | 85.7 | 104 | 8 | ADS22888 | Ads22888 Bacterial |
| 12 | 76 | 83.5 | 19 | 4 | AAU04219 | Aau04219 Chaperoni |
| 13 | 76 | 83.5 | 19 | 4 | AAU04226 | Aau04226 Chaperoni |
| 14 | 76 | 83.5 | 98 | 8 | ADS23298 | Ads23298 Bacterial |
| 15 | 76 | 83.5 | 122 | 8 | ADN17513 | Adn17513 Bacterial |
| 16 | 75 | 82.4 | 19 | 4 | AAU04233 | Aau04233 Chaperoni |
| 17 | 75 | 82.4 | 101 | 2 | AAW16679 | Aaw16679 Lawsonia |
| 18 | 75 | 82.4 | 104 | 2 | ADS42281 | Ads42281 Bacterial |
| 19 | 73 | 80.2 | 19 | 4 | AAU04240 | Aau04240 Chaperoni |
| 20 | 73 | 80.2 | 51 | 2 | AAW89986 | Aaw89986 Expressed |
| 21 | 73 | 80.2 | 72 | 2 | AAW89858 | Aaw89858 Antigen 1 |
| 22 | 73 | 80.2 | 100 | 8 | ADP03527 | Adp03527 G. entani |
| 23 | 73 | 80.2 | 104 | 5 | ABU52127 | Abu52127 Helicobac |
| 24 | 73 | 80.2 | 118 | 5 | AAR74338 | Aar74338 Helicobac |
| 25 | 73 | 80.2 | 118 | 2 | AAR67373 | Aar67373 Heat shoc |
| 26 | 73 | 80.2 | 118 | 2 | AAW06731 | Aaw06731 H. pylori |
| 27 | 73 | 80.2 | 118 | 4 | AAU35688 | Aau35688 Helicobac |
| 28 | 73 | 80.2 | 118 | 4 | AAU35872 | Aau35872 Helicobac |
| 29 | 73 | 80.2 | 118 | 7 | ADG75218 | Adg75218 Helicobac |
| 30 | 73 | 80.2 | 118 | 7 | ADG75220 | Adg75220 Helicobac |
| 31 | 72 | 79.1 | 89 | 6 | ADS21261 | Ads21261 Bacterial |
| 32 | 72 | 79.1 | 96 | 6 | ABP77243 | Abp77243 N. gonorr |
| 33 | 72 | 79.1 | 96 | 8 | ADS20982 | Ads20982 Bacterial |
| 34 | 72 | 79.1 | 99 | 4 | AAB78872 | Aab78872 C. glutam |
| 35 | 72 | 79.1 | 99 | 4 | AAG90410 | Aag90410 C. glutami |
| 36 | 72 | 79.1 | 19 | 4 | ABM42247 | Abm42247 Propionib |
| 37 | 71 | 78.0 | 130 | 4 | AAU45728 | Aau45728 Propionib |
| 38 | 71 | 78.0 | 19 | 4 | AAU04222 | Aau04222 Chaperoni |
| 39 | 71 | 78.0 | 95 | 8 | ADS43337 | Ads43337 Bacterial |
| 40 | 71 | 78.0 | 92 | 8 | ADN20296 | Adn20296 Bacterial |
| 41 | 70 | 76.9 | 98 | 8 | ADS25886 | Ads25886 Bacterial |
| 42 | 70 | 76.9 | 98 | 8 | ADS26231 | Ads26231 Bacterial |
| 43 | 70 | 76.9 | 98 | 8 | ADS26231 | Ads26231 Bacterial |
| 44 | 70 | 76.9 | 98 | 8 | ADS22726 | Ads22726 Bacterial |
| 45 | 70 | 76.9 | 98 | 8 | ADS25173 | Ads25173 Bacterial |

## ALIGNMENTS

RESULT 1
AAU04230

AAU04230 standard; peptide; 19 AA.

XX
ID   AAU04230 standard; peptide; 19 AA.
XX
AC   AAU04230;
XX
DT   23-OCT-2001 (first entry)
XX
DE   Chaperonin binding domain #19.
XX
KW   Chaperonin binding domain; therapeutic; protein expression;
KW   growth factor; cytokine; ligand; receptor; inhibitor; vaccine;
KW   industrial protein.
XX
OS   Unidentified.
XX
PN   WO200148227-A1.
XX
PD   05-JUL-2001.
XX
PF   14-DEC-2000; 2000WO-US034055.
XX
PR   23-DEC-1999; 99US-00470830.
XX
PA   (GEMV ) GENENCOR INT INC.
XX
PI   Joachimiak A, Donelly M;
XX
DR   WPI; 2001-425674/45.
XX
PT   Producing proteins, particularly toxic proteins in host cells, such as
PT   bacterial cells, by using a chaperonin binding domain in expression
PT   systems designed for the production of the proteins.
XX
PS   Claim 9; Page 8; 26pp; English.
XX
CC   The sequence represents the amino acid sequence of chaperonin binding
CC   domain #19. The sequence is used in a method of producing a protein in a
CC   host cell, comprising culturing a host cell comprising a nucleic acid
CC   encoding an isolated chaperonin binding domain (CBD), associated with a
CC   second nucleic acid encoding the protein and a third nucleic acid
CC   encoding a chaperonin. This is carried out under conditions suitable for
CC   expression of the nucleic acids, where CBD is capable of binding to the
CC   chaperonin. The method is useful for producing a protein, particularly
CC   toxic protein in a bacterial cell, especially Escherichia coli. CBD is
CC   useful for producing heterologous proteins, peptides or polypeptides in a
CC   host cell, including therapeutically significant proteins, such as growth
CC   factors, cytokines, ligands, receptors and inhibitors, vaccines,

OM protein - protein search, using sw model

           GenCore version 5.1.6
           Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:       September 29, 2005, 19:34:24 ; Search time 24.5 Seconds
                                    (without alignments)
                                    74.617 Million cell updates/sec

Title:       US-10-714-212-21
Perfect score:  91
Sequence:     1 ESEEKTASGIVLPDSAKEK 19
Scoring table:  BLOSUM62
              Gapop 10.0 , Gapext 0.5

Searched:     283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:     283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      PIR.79:*
            1: pir1:*
            2: pir2:*
            3: pir3:*
            4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | %Query Match Length | DB | ID | Description |
|---|---|---|---|---|---|
| 1 | 100.0 | 94 | 2 | A41884 | heat shock protein |
| 2 | 96.7 | 94 | 2 | A81333 | class I heat-shock |
| 3 | 96.7 | 94 | 2 | AD1704 | class I heat-shock |
| 4 | 93.4 | 88 | 2 | PC4238 | heat shock protein |
| 5 | 93.4 | 94 | 2 | JC1479 | heat shock protein |
| 6 | 93.4 | 94 | 2 | A49855 | heat shock protein |
| 7 | 93.4 | 94 | 2 | A83720 | class I heat-shock |
| 8 | 87.9 | 103 | 1 | BVYCGS | chaperonin groES |
| 9 | 80.0 | 103 | 2 | A36721 | groES protein - Sy |
| 10 | 83.5 | 98 | 2 | JN0510 | heat shock protein |
| 11 | 83.5 | 101 | 2 | JN0513 | heat shock protein |
| 12 | 76.1 | 122 | 2 | B70489 | GroES - Aquifex ae |
| 13 | 76.2 | 98 | 1 | A43827 | chaperonin groES - |
| 14 | 82.4 | 98 | 2 | D95311 | chaperonin groES - |
| 15 | 82.4 | 104 | 2 | AF3640 | 10K chaperonin gro |
| 16 | 82.4 | 103 | 2 | S35310 | groES2 chaperonin |
| 17 | 81.3 | 103 | 2 | T06830 | heat shock cognate |
| 18 | 80.2 | 95 | 2 | B71668 | chaperonin groES - |
| 19 | 80.2 | 106 | 2 | S77563 | 10 kd chaperonin ( |
| 20 | 73.0 | 118 | 2 | C71986 | 10kda chaperone - |
| 21 | 80.2 | 118 | 2 | S61396 | chaperonin groES - |
| 22 | 79.1 | 95 | 2 | I40330 | Cpn10 protein (Gro |
| 23 | 79.1 | 95 | 2 | G81019 | chaperonin, 10 kDa |
| 24 | 72.1 | 96 | 2 | GB1964 | chaperonin 10 Kd B |
| 25 | 79.1 | 103 | 2 | AF2263 | chaperonin GroES | |
| 26 | 71.1 | 88 | 2 | A71253 | probable chaperoni |
| 27 | 78.0 | 95 | 2 | A41872 | heat shock protein |
| 28 | 78.0 | 95 | 2 | G9232 | co-chaperonin GroE |
| 29 | 71.1 | 104 | 2 | S35308 | heat shock protein |
| 30 | 78.0 | 71 | 2 | F75499 | chaperonin - Deino |
| 31 | 76.9 | 92 | 2 | G72367 | groES protein - Th |
| 32 | 76.9 | 94 | 2 | S68248 | chaperonin groES - |
| 33 | 76.9 | 98 | 2 | C97442 | 10K chaperonin (pr |
| 34 | 76.9 | 98 | 2 | A22660 | 10 KD chaperonin ( |
| 35 | 76.9 | 98 | 2 | A36917 | heat shock protein |
| 36 | 75.8 | 100 | 1 | BVMY7B | chaperonin groES - |
| 37 | 75.8 | 100 | 1 | BVMYBA | chaperonin groES - |
| 38 | 75.8 | 102 | 2 | A41325 | heat shock protein |
| 39 | 75.8 | 102 | 2 | S37565 | chaperonin groES - |
| 40 | 74.7 | 95 | 2 | JC2563 | chaperonin GroES [ |
| 41 | 74.7 | 99 | 2 | A97821 | 10K chaperonin [im |
| 42 | 72.5 | 96 | 2 | C87334 | chaperonin, 10 kDa |
| 43 | 72.5 | 99 | 2 | S70668 | chaperonin groES ( |
| 44 | 71.4 | 86 | 2 | F81328 | 10 kD chaperonin |
| 45 | 71.4 | 94 | 2 | JN0600 | heat shock protein |

ALIGNMENTS

RESULT 1
A41884
heat shock protein (chaperonin) groES - Bacillus subtilis
N;Alternate names: class I heat-shock protein
C;Species: Bacillus subtilis
C;Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A41884; A41885; JC1371; E69637
R;Li, M.; Wong, S.L.
J. Bacteriol. 174, 3981-3992, 1992
A;Title: Cloning and characterization of the groESL operon from Bacillus subtilis.
A;Reference number: A41884; MUID:92283753; PMID:1350776
A;Accession: A41884
A;Molecule type: DNA
A;Residues: 1-94 <LI1>
A;Cross-references: UNIPROT:P28599; GB:M81132; NID:g143025; PIDN:AAA22502.1; PID:g14302
A;Experimental source: 168, trpC2
A;Note: this sequence is inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:105982, NCBIP:105986)
R;Schmidt, A.; Schiesswohl, M.; Volker, U.; Hecker, M.; Schumann, W.
J. Bacteriol. 174, 3993-3999, 1992
A;Title: Cloning, sequencing, mapping, and transcriptional analysis of the groESL opero
A;Reference number: A41885; MUID:92283754; PMID:1350777
A;Accession: A41885
A;Molecule type: DNA
A;Residues: 1-94 <SCH>
A;Cross-references: GB:M84965; NID:g143061; PIDN:AAA22530.1; PID:g143062
A;Experimental source: MB11
A;Note: sequence extracted from NCBI backbone (NCBIN:105432, NCBIP:105433)
R;Tozawa, Y.; Yoshikawa, H.; Kawamura, F.; Itaya, M.; Takahashi, H.
Biosci. Biotechnol. Biochem. 56, 1995-2002, 1992
A;Title: Isolation and characterization of the groESL and groEL genes of Bacillus subtil
A;Reference number: JC1371; MUID:93129852; PMID:1369494
A;Accession: JC1371
A;Molecule type: DNA
A;Residues: 1-94 <TOZ>
A;Cross-references: GB:D10972; GB:D01157; NID:g4433779; PIDN:BAA22518.1; PID:g4433780
A;Experimental source: strain Marburg 168
A;Note: the authors translated the codon ACT for residue 20 as Tyr
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
ich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama

OM protein - protein search, using sw model

Copyright (c) 1993 - 2005 Compugen Ltd.
GenCore version 5.1.6

Run on:            September 29, 2005, 19:33:34 ; Search time 113.5 Seconds
                   (without alignments)
                   85.723 Million cell updates/sec

Title:             US-10-714-212-21
Perfect score:     91
Sequence:          1 ESEEKTASGIVLPDSAKEK 19

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:     1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing:   Minimum Match 0%
                   Maximum Match 100%
                   Listing first 45 summaries

Database :         UniProt_03:*
                   1: uniprot_sprot:*
                   2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 91 | 100.0 | 94 | 1 | CH10_BACSU | P28599 bacillus su |
| 2 | 91 | 100.0 | 94 | 2 | Q65MZ9 | Q65mz9 bacillus li |
| 3 | 88 | 96.7 | 94 | 1 | CH10_LISIN | Q929u9 listeria in |
| 4 | 88 | 96.7 | 94 | 1 | CH10_LISMO | Q9age7 listeria mo |
| 5 | 88 | 96.7 | 94 | 2 | Q71XU5 | Q71xu5 listeria mo |
| 6 | 88 | 96.7 | 94 | 1 | CH10_BACHD | O50304 bacillus ha |
| 7 | 85 | 93.4 | 94 | 2 | Q6HBL3 | P26210 bacillus ps |
| 8 | 85 | 93.4 | 94 | 1 | CH10_BACP3 | O07200 bacillus st |
| 9 | 85 | 93.4 | 94 | 1 | CH10_BACST | Q8vv85 bacillus th |
| 10 | 85 | 93.4 | 94 | 1 | CH10_BACTR | Q7djl2 bacillus th |
| 11 | 85 | 93.4 | 94 | 2 | Q8VV85 | bacillus sp |
| 12 | 85 | 92.3 | 94 | 2 | Q7DJL2 | bacillus st |
| 13 | 84 | 92.3 | 103 | 1 | Q9EZV5 | bacillus sp |
| 14 | 82 | 90.1 | 94 | 2 | Q7TU43 | prochloroco |
| 15 | 82 | 90.1 | 94 | 2 | Q63GV8 | bacillus ce |
| 16 | 82 | 90.1 | 94 | 2 | Q73ES0 | bacillus st |
| 17 | 82 | 90.1 | 96 | 2 | Q63GV8 | prochloroco |
| 18 | 81 | 90.1 | 93 | 1 | Q81VE2 | bacillus ce |
| 19 | 81 | 89.0 | 94 | 2 | Q6HPC8 | bacillus ce |
| 20 | 81 | 89.0 | 94 | 2 | Q81VE2 | bacillus an |
| 21 | 81 | 89.0 | 94 | 2 | Q814B1 | bacillus th |
| 22 | 80 | 87.9 | 103 | 1 | Q8CX14 | bacillus ce |
| 23 | 80 | 87.9 | 103 | 1 | Q7TTX2 | oceanobacil |
| 24 | 78 | 85.7 | 101 | 1 | Q7TUS3 | synechococc |
| 25 | 78 | 85.7 | 103 | 2 | Q7TTX2 | synechococc |
| 26 | 78 | 85.7 | 104 | 1 | Q7TV92 | prochloroco |
| 27 | 78 | 85.7 | 104 | 2 | P07889 | prochloroco |
| 28 | 78 | 85.7 | 89 | 1 | P22880 | synechococc |
| 29 | 76 | 83.5 | 104 | 1 | Q7UM98 | synechococc |
| 30 | 76 | 83.5 | 97 | 2 | Q7UN98 | rhodopirell |
| 31 | 76 | 83.5 | 98 | 1 | CH11_RHIME | rhizobium 1 |
| 32 | 76 | 83.5 | 98 | 1 | CH13_RHILO | Q98ax8 rhizobium 1 |
| 33 | 76 | 83.5 | 98 | 1 | CH14_RHILO | Q983s3 rhizobium 1 |
| 34 | 76 | 83.5 | 98 | 1 | CH15_RHILO | Q981k0 rhizobium 1 |
| 35 | 76 | 83.5 | 98 | 2 | Q799Q4 | Q799q4 rhizobium 1 |
| 36 | 76 | 83.5 | 98 | 2 | Q6FYU9 | Q6fyu9 barconella |
| 37 | 76 | 83.5 | 98 | 2 | Q6H8L3 | Q6h813 barconella |
| 38 | 76 | 83.5 | 103 | 1 | CH10_SYNP8 | O81373 synechococc |
| 39 | 76 | 83.5 | 104 | 1 | CH15_RHIME | P35474 rhizobium m |
| 40 | 76 | 82.4 | 122 | 1 | CH10_AQUAE | O67942 aquifex aeo |
| 41 | 76 | 82.4 | 93 | 2 | Q7TTZ0 | O7ttz0 rhodopirell |
| 42 | 75 | 82.4 | 95 | 2 | Q72AL5 | O72a15 desulfotale |
| 43 | 75 | 82.4 | 95 | 1 | Q6ARV5 | Q6arv5 desulfovibr |
| 44 | 75 | 82.4 | 98 | 1 | CH10_BRUME | P25968 brucella me |
| 45 | 75 | 82.4 | 98 | 1 | CH14_RHIME | Q92zq3 rhizobium m |

## ALIGNMENTS

RESULT 1
CH10_BACSU

```
ID   CH10_BACSU   STANDARD;   PRT;   94 AA.
AC   P28599; O08341;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   10 kDa chaperonin (Protein Cpn10) (groES protein).
GN   Name=groS; Synonyms=groES, mopB; OrderedLocusNames=BSU06020;
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=92283753; PubMed=1350776;
RA   Li M., Wong S.L.;
RT   "Cloning and characterization of the groESL operon from Bacillus
RT   subtilis.";
RL   J. Bacteriol. 174:3981-3992(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92283754; PubMed=1350777;
RA   Schmidt A., Schiesswohl M., Voelker U., Hecker M., Schumann W.;
RT   "Cloning, sequencing, mapping, and transcriptional analysis of the
RT   groESL operon from Bacillus subtilis.";
RL   J. Bacteriol. 174:3993-3999(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / Marburg;
RX   MEDLINE=93129852; PubMed=1369494;
RA   Tozawa Y., Yoshikawa H., Kawamura F., Itaya M., Takahashi H.;
RT   "Isolation and characterization of the groES and groEL genes of
RT   Bacillus subtilis Marburg.";
RL   Biosci. Biotechnol. Biochem. 56:1995-2002(1992).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / JH642;
RX   MEDLINE=97346038; PubMed=9202461;
RA   Sadaie Y., Yata K., Fujita M., Sagai H., Itaya M., Kasahara Y.,
RA   Ogasawara N.;
RT   "Nucleotide sequence and analysis of the phoB-rrnE-groESL region of
RT   the Bacillus subtilis chromosome.";
RL   Microbiology 143:1861-1866(1997).
RN   [5]
RP   SEQUENCE OF 5-94 FROM N.A.
RC   STRAIN=168 / Marburg;
RX   MEDLINE=98116660; PubMed=9455482;
RA   Kasahara Y., Nakai S., Ogasawara N., Yata K., Sadaie Y.;
RT   "Sequence analysis of the groESL-cota region of the Bacillus subtilis
RT   genome, containing the restriction/modification system genes.";
RL   DNA Res. 4:335-339(1997).
RN   [6]
RP   SEQUENCE FROM N.A.
```

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:          September 29, 2005, 19:35:05 ; Search time 30.5 Seconds
                                               (without alignment)
                                               46.503 Million cell updates/sec

Title:           US-10-714-212-21
Perfect score:   91
Sequence:        1 ESEEKTASGIVLPDSAKEK 19

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:    513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | % Query Match | Length | DB ID | Description |
|---|---|---|---|---|---|
| 1 | 100.0 | 19 | 4 | US-09-470-830A-21 | Sequence 21, Appl |
| 2 | 85 | 93.4 | 19 | 4 | US-09-470-830A-20 | Sequence 20, Appl |
| 3 | 85 | 93.4 | 19 | 4 | US-09-470-830A-22 | Sequence 22, Appl |
| 4 | 80 | 87.9 | 19 | 4 | US-09-470-830A-29 | Sequence 29, Appl |
| 5 | 80 | 87.9 | 19 | 4 | US-09-470-830A-30 | Sequence 30, Appl |
| 6 | 80 | 87.9 | 19 | 2 | US-08-467-822-38 | Sequence 38, Appl |
| 7 | 80 | 87.9 | 94 | 3 | US-08-432-697-38 | Sequence 38, Appl |
| 8 | 80 | 87.9 | 94 | 3 | US-08-432-697-38 | Sequence 38, Appl |
| 9 | 80 | 87.9 | 94 | 3 | US-08-466-248-38 | Sequence 38, Appl |
| 10 | 76 | 83.5 | 19 | 4 | US-09-470-830A-10 | Sequence 10, Appl |
| 11 | 76 | 83.5 | 19 | 4 | US-09-470-830A-17 | Sequence 17, Appl |
| 12 | 75 | 82.4 | 19 | 4 | US-09-470-830A-12 | Sequence 12, Appl |
| 13 | 75 | 82.4 | 19 | 4 | US-09-470-830A-24 | Sequence 24, Appl |
| 14 | 73 | 80.2 | 19 | 4 | US-09-470-830A-31 | Sequence 31, Appl |
| 15 | 73 | 80.2 | 118 | 3 | US-08-467-822-29 | Sequence 29, Appl |
| 16 | 73 | 80.2 | 118 | 3 | US-08-466-248-29 | Sequence 29, Appl |
| 17 | 72 | 79.1 | 99 | 4 | US-09-603-203A-4 | Sequence 4, Appli |
| 18 | 71 | 78.0 | 19 | 4 | US-09-470-830A-13 | Sequence 13, Appl |
| 19 | 71 | 78.0 | 19 | 4 | US-09-470-830A-15 | Sequence 15, Appl |
| 20 | 70 | 76.9 | 19 | 4 | US-09-470-830A-19 | Sequence 19, Appl |
| 21 | 69 | 75.8 | 19 | 4 | US-09-470-830A-23 | Sequence 23, Appl |
| 22 | 69 | 75.8 | 19 | 4 | US-09-470-830A-26 | Sequence 26, Appl |
| 23 | 69 | 75.8 | 198 | 4 | US-09-613-303-35 | Sequence 35, Appl |
| 24 | 69 | 75.8 | 198 | 4 | US-10-267-311-35 | Sequence 35, Appl |
| 25 | 65 | 71.4 | 19 | 4 | US-09-470-830A-25 | Sequence 25, Appl |
| 26 | 65 | 71.4 | 198 | 4 | US-09-134-001C-4686 | Sequence 4686, Ap |
| 27 | 65 | 71.4 | 111 | 3 | | Sequence 4686, Ap |
| 28 | 64 | 70.3 | 19 | 4 | US-09-470-830A-6 | Sequence 6, Appli |
| 29 | 64 | 70.3 | 19 | 4 | US-09-470-830A-9 | Sequence 9, Appli |
| 30 | 64 | 70.3 | 19 | 4 | US-09-470-830A-11 | Sequence 11, Appl |
| 31 | 64 | 70.3 | 94 | 2 | US-08-467-822-37 | Sequence 37, Appl |
| 32 | 64 | 70.3 | 94 | 3 | US-08-432-697-37 | Sequence 37, Appl |
| 33 | 64 | 70.3 | 94 | 3 | US-08-466-248-37 | Sequence 37, Appl |
| 34 | 64 | 70.3 | 140 | 4 | US-09-540-236-2611 | Sequence 2611, Ap |
| 35 | 64 | 70.3 | 145 | 4 | US-09-252-991A-24284 | Sequence 24284, A |
| 36 | 61 | 67.0 | 106 | 4 | US-09-328-352-5510 | Sequence 5510, Ap |
| 37 | 61 | 67.0 | 19 | 4 | US-09-470-830A-8 | Sequence 8, Appli |
| 38 | 60 | 65.9 | 19 | 4 | US-09-470-830A-27 | Sequence 27, Appl |
| 39 | 60 | 65.9 | 19 | 4 | US-09-470-830A-7 | Sequence 7, Appli |
| 40 | 59 | 64.8 | 93 | 2 | US-08-467-822-36 | Sequence 36, Appl |
| 41 | 59 | 64.8 | 93 | 3 | US-08-432-697-36 | Sequence 36, Appl |
| 42 | 59 | 64.8 | 93 | 3 | US-08-466-248-36 | Sequence 36, Appl |
| 43 | 58.5 | 64.3 | 102 | 4 | US-09-439-185A-137 | Sequence 137, App |
| 44 | 58 | 63.7 | 119 | 4 | US-09-107-532A-6969 | Sequence 6969, Ap |
| 45 | 57.5 | 63.2 | 102 | 4 | US-09-198-452A-154 | Sequence 154, App |

## ALIGNMENTS

RESULT 1
US-09-470-830A-21

; Sequence 21, Application US/0947083OA
; Patent No. 6677139
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; TITLE OF INVENTION: Host Cells
; FILE REFERENCE: GC559-PCT
; CURRENT APPLICATION NUMBER: US/09/470,830A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/470,830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-470-830A-21

  Query Match             100.0%;  Score 91;  DB 4;  Length 19;
  Best Local Similarity   100.0%;
  Matches    19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ESEEKTASGIVLPDSAKEK 19
              |||||||||||||||||||
Db          1 ESEEKTASGIVLPDSAKEK 19

RESULT 2
US-09-470-830A-20

; Sequence 20, Application US/0947083OA
; Patent No. 6677139
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; TITLE OF INVENTION: Host Cells
; FILE REFERENCE: GC559-PCT
; CURRENT APPLICATION NUMBER: US/09/470,830A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/470,830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Thermoactinomyces sp.
US-09-470-830A-20

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 29, 2005, 19:38:05 ; Search time 117 Seconds
(without alignments)
67.412 Million cell updates/sec

Title:          US-10-714-212-21
Perfect score:  91
Sequence:       1 ESEEKTASGIVLPDSAKEK 19
Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched:       1846076 seqs, 415516000 residues
Total number of hits satisfying chosen parameters:   1846076
Post-processing: Minimum Match 0%
                 Maximum Match 100%
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                 Listing first 45 summaries

Database :  Published Applications AA:*
  1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
  2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
  3: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
  4: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
  5: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
  6: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
  7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
  8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
  9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
 12: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
 13: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
 19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
 20: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % | Query Match Length | DB ID | | Description |
|---|---|---|---|---|---|---|
| 1 | 91 | 100.0 | 19 | 15 | US-10-714-212-21 | Sequence 21, Appl |
| 2 | 85 | 93.4 | 19 | 15 | US-10-714-212-20 | Sequence 20, Appl |
| 3 | 85 | 93.4 | 19 | 15 | US-10-714-212-22 | Sequence 22, Appl |
| 4 | 81 | 89.0 | 94 | 15 | US-10-369-493-10863 | Sequence 10863, A |
| 5 | 80 | 87.9 | 19 | 15 | US-10-714-212-29 | Sequence 29, Appl |
| 6 | 80 | 87.9 | 19 | 15 | US-10-714-212-30 | Sequence 30, Appl |
| 7 | 78 | 85.7 | 104 | 15 | US-10-369-493-17908 | Sequence 17908, A |
| 8 | 78 | 85.7 | 95 | 15 | US-10-369-493-11921 | Sequence 11921, A |
| 9 | 76 | 83.5 | 19 | 15 | US-10-714-212-10 | Sequence 10, Appl |
| 10 | 76 | 83.5 | 19 | 15 | US-10-714-212-17 | Sequence 17, Appl |
| 11 | 76 | 83.5 | 98 | 15 | US-10-369-493-12331 | Sequence 12331, A |
| 12 | 76 | 83.5 | 122 | 15 | US-10-369-493-166 | Sequence 166, Appl |
| 13 | 75 | 82.4 | 19 | 15 | US-10-714-212-12 | Sequence 12, Appl |
| 14 | 75 | 82.4 | 19 | 15 | US-10-714-212-24 | Sequence 24, Appl |
| 15 | 75 | 82.4 | 104 | 15 | US-10-369-493-20711 | Sequence 20711, A |
| 16 | 75 | 82.4 | 19 | 15 | US-10-714-212-31 | Sequence 31, Appl |
| 17 | 73 | 80.2 | 118 | 9 | US-09-815-242-11281 | Sequence 11281, A |
| 18 | 73 | 80.2 | 118 | 9 | US-09-815-242-11465 | Sequence 11465, A |
| 19 | 73 | 80.2 | 118 | 18 | US-10-953-901-419 | Sequence 419, App |
| 20 | 73 | 80.2 | 118 | 18 | US-10-953-901-421 | Sequence 421, App |
| 21 | 72 | 79.1 | 96 | 15 | US-10-369-493-10294 | Sequence 10294, A |
| 22 | 72 | 79.1 | 99 | 9 | US-09-738-626-4164 | Sequence 4164, Ap |
| 23 | 72 | 79.1 | 19 | 15 | US-10-369-493-10015 | Sequence 10015, A |
| 24 | 71 | 78.0 | 19 | 15 | US-10-714-212-13 | Sequence 13, Appl |
| 25 | 71 | 78.0 | 19 | 15 | US-10-714-212-15 | Sequence 15, Appl |
| 26 | 71 | 78.0 | 95 | 15 | US-10-369-493-21767 | Sequence 21767, A |
| 27 | 70 | 76.9 | 19 | 15 | US-10-714-212-14 | Sequence 14, Appl |
| 28 | 70 | 76.9 | 92 | 15 | US-10-369-493-11759 | Sequence 11759, A |
| 29 | 70 | 76.9 | 98 | 15 | US-10-369-493-2949 | Sequence 2949, Ap |
| 30 | 70 | 76.9 | 98 | 15 | US-10-369-493-11759 | Sequence 11759, A |
| 31 | 70 | 76.9 | 98 | 15 | US-10-369-493-14206 | Sequence 14206, A |
| 32 | 70 | 76.9 | 98 | 15 | US-10-369-493-14459 | Sequence 14459, A |
| 33 | 70 | 76.9 | 98 | 15 | US-10-369-493-14919 | Sequence 14919, A |
| 34 | 70 | 76.9 | 98 | 15 | US-10-369-493-15264 | Sequence 15264, A |
| 35 | 69 | 75.8 | 19 | 15 | US-10-714-212-19 | Sequence 19, Appl |
| 36 | 69 | 75.8 | 19 | 15 | US-10-369-493-7916 | Sequence 23, Appl |
| 37 | 69 | 75.8 | 100 | 9 | US-09-712-363-272 | Sequence 272, App |
| 38 | 69 | 75.8 | 100 | 16 | US-10-332-512A-11 | Sequence 11, Appl |
| 39 | 69 | 75.8 | 102 | 14 | US-10-156-761-12525 | Sequence 12525, A |
| 40 | 69 | 75.8 | 198 | 14 | US-10-267-311-35 | Sequence 35, Appl |
| 41 | 69 | 75.8 | 198 | 17 | US-10-677-956-35 | Sequence 35, Appl |
| 42 | 68 | 74.7 | 95 | 15 | US-10-369-493-7916 | Sequence 7916, Ap |
| 43 | 67 | 73.6 | 101 | 15 | US-10-369-493-20609 | Sequence 20609, A |
| 44 | 66 | 72.5 | 96 | 15 | US-10-369-493-16818 | Sequence 16818, A |
| 45 | 66 | 72.5 | 97 | 15 | US-10-369-493-14074 | Sequence 14074, A |

ALIGNMENTS

RESULT 1
US-10-714-212-21
; Sequence 21, Application US/10714212
; Publication No. US20040077038A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; TITLE OF INVENTION: Host Cells
; FILE REFERENCE: GC559-PCT
; CURRENT APPLICATION NUMBER: US-10/714,212
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 09/470,830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-714-212-21

Query Match       100.0%;  Score 91;  DB 15;  Length 19;
Best Local Similarity   100.0%;
Matches   19;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 ESEEKTASGIVLPDSAKEK 19
             |||||||||||||||||||
Db         1 ESEEKTASGIVLPDSAKEK 19

RESULT 2
US-10-714-212-20
; Sequence 20, Application US/10714212
; Publication No. US20040077038A1

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        September 29, 2005, 19:30:39 ; Search time 118 Seconds.
               (without alignments)
               62.275 Million cell updates/sec

Title:         US-10-714-212-22
Perfect score: 93
Sequence:      1 ETEEKTASGIVLPDTAKEK 19
Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:        2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    A_Geneseq_16Dec04:*
         1:   geneseqp1980s:*
         2:   geneseqp1990s:*
         3:   geneseqp2000s:*
         4:   geneseqp2001s:*
         5:   geneseqp2002s:*
         6:   geneseqp2003as:*
         7:   geneseqp2003bs:*
         8:   geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 93 | 100.0 | 19 | 4 | AAU04229 | Aau04229 Chaperoni |
| 2 | 93 | 100.0 | 19 | 4 | AAU04231 | Aau04231 Chaperoni |
| 3 | 93 | 100.0 | 40 | 2 | AAR20197 | Aar20197 Heat shoc |
| 4 | 88 | 94.6 | 94 | 2 | AAR67388 | Aar67388 GroES-lik |
| 5 | 85 | 91.4 | 19 | 4 | AAU04230 | Aau04230 Chaperoni |
| 6 | 84 | 90.3 | 94 | 5 | ABB49470 | Abb49470 Listeria |
| 7 | 81 | 87.1 | 95 | 8 | ADS28875 | Ads28875 Bacterial |
| 8 | 81 | 87.1 | 104 | 8 | ADS22888 | Ads22888 Bacterial |
| 9 | 79 | 84.9 | 122 | 8 | ADN17513 | Adn17513 Bacterial |
| 10 | 78 | 83.9 | 19 | 4 | AAU04238 | Aau04238 Chaperoni |
| 11 | 78 | 83.9 | 19 | 4 | AAU04233 | Aau04233 Chaperoni |
| 12 | 78 | 83.9 | 19 | 4 | AAU04239 | Aau04239 Chaperoni |
| 13 | 78 | 83.9 | 104 | 4 | ADS42281 | Ads42281 Bacterial |
| 14 | 77 | 82.8 | 19 | 4 | AAU04219 | Aau04219 Chaperoni |
| 15 | 77 | 82.8 | 19 | 4 | AAU04226 | Aau04226 Chaperoni |
| 16 | 77 | 82.8 | 94 | 8 | ADS21830 | Ads21830 Bacterial |
| 17 | 77 | 82.8 | 98 | 8 | ADS23298 | Ads23298 Bacterial |
| 18 | 76 | 81.7 | 101 | 2 | AAW16679 | Aaw16679 Lawsonia |
| 19 | 75 | 80.6 | 89 | 8 | ADS21261 | Ads21261 Bacterial |
| 20 | 75 | 80.6 | 130 | 4 | AAU45728 | Aau45728 Propionib |
| 21 | 75 | 80.6 | 130 | 6 | ABM42247 | Abm42247 Propionib |
| 22 | 74 | 79.6 | 19 | 4 | AAU04222 | Aau04222 Chaperoni |
| 23 | 74 | 79.6 | 100 | 8 | ADP03527 | Adp03527 G. entani |
| 24 | 73 | 78.5 | 96 | 8 | ADS20982 | Ads20982 Bacterial |
| 25 | 72 | 77.4 | 19 | 4 | AAU04232 | Aau04232 Chaperoni |
| 26 | 72 | 77.4 | 19 | 4 | AAU04228 | Aau04228 Chaperoni |
| 27 | 72 | 77.4 | 95 | 8 | ADS43337 | Ads43337 Bacterial |
| 28 | 72 | 77.4 | 99 | 2 | AAR05700 | Aar05700 MPB-57 pr |
| 29 | 72 | 77.4 | 100 | 4 | AAG81221 | Aag81221 Mycobacte |
| 30 | 72 | 77.4 | 100 | 5 | AAM50739 | Aam50739 Mycobacte |
| 31 | 72 | 77.4 | 100 | 8 | AAE24631 | Aae24631 Mycobacte |
| 32 | 72 | 77.4 | 102 | 2 | ADN11339 | Adn11339 Chaperoni |
| 33 | 72 | 77.4 | 198 | 2 | AAR22364 | Aar22364 GroES str |
| 34 | 71 | 76.3 | 19 | 4 | AAB31616 | Aab31616 Amino aci |
| 35 | 71 | 76.3 | 19 | 4 | AAU04240 | Aau04240 Chaperoni |
| 36 | 71 | 76.3 | 72 | 2 | AAW89858 | Aaw89858 Antigen 1 |
| 37 | 71 | 76.3 | 51 | 2 | AAW89986 | Aaw89986 Expressed |
| 38 | 71 | 76.3 | 98 | 8 | ADS25886 | Ads25886 Bacterial |
| 39 | 71 | 76.3 | 98 | 8 | ADS26231 | Ads26231 Bacterial |
| 40 | 71 | 76.3 | 98 | 8 | ADS22726 | Ads22726 Bacterial |
| 41 | 71 | 76.3 | 98 | 8 | ADS25173 | Ads25173 Bacterial |
| 42 | 71 | 76.3 | 98 | 8 | ADS25426 | Ads25426 Bacterial |
| 43 | 71 | 76.3 | 104 | 8 | ABU52127 | Abu52127 Helicobac |
| 44 | 71 | 76.3 | 118 | 2 | AAR74338 | Aar74338 Helicobac |
| 45 | 71 | 76.3 | 118 | 2 | AAR67373 | Aar67373 Heat shoc |

ALIGNMENTS

RESULT 1
AAU04229
AAU04229  AAU04229 standard; peptide; 19 AA.
XX
ID  AAU04229;
XX
AC  AAU04229;
XX
DT  23-OCT-2001  (first entry)
XX
DE  Chaperonin binding domain #18.
XX
KW  Chaperonin binding domain; therapeutic; protein expression;
KW  growth factor; cytokine; ligand; receptor; inhibitor; vaccine;
KW  industrial protein.
XX
OS  Unidentified.
XX
PN  WO200148227-A1.
XX
PD  05-JUL-2001.
XX
PF  14-DEC-2000; 2000WO-US034055.
XX
PR  23-DEC-1999;  99US-00470830.
XX
PA  (GEMV ) GENENCOR INT INC.
XX
PI  Joachimiak A,  Donelly M;
XX
DR  WPI; 2001-425674/45.
XX
PT  Producing proteins, particularly toxic proteins in host cells, such as
PT  bacterial cells, by using a chaperonin binding domain in expression
PT  systems designed for the production of the proteins.
XX
PS  Claim 9; Page 8; 26pp; English.
XX
CC  The sequence represents the amino acid sequence of chaperonin binding
CC  domain #18. The sequence is used in a method of producing a protein in a
CC  host cell, comprising culturing a host cell comprising a nucleic acid
CC  encoding an isolated chaperonin binding domain (CBD), associated with a
CC  second nucleic acid encoding the protein and a third nucleic acid
CC  encoding a chaperonin. This is carried out under conditions suitable for
CC  expression of the nucleic acids, where CBD is capable of binding to the
CC  chaperonin. The method is useful for producing a protein, particularly
CC  toxic protein in a bacterial cell, especially Escherichia coli. CBD is
CC  useful for producing heterologous proteins, peptides or polypeptides in a
CC  host cell, including therapeutically significant proteins, such as growth
CC  factors, cytokines, ligands, receptors and inhibitors, vaccines,

OM protein - protein search, using sw model

Run on:    September 29, 2005, 19:34:24 ; Search time 24.5 Seconds
              (without alignments)
              74.617 Million cell updates/sec

Title:        US-10-714-212-22
Perfect score:  93
Sequence:     1 ETEEKTASGIVLPDTAKEK 19

Scoring table:  BLOSUM62
              Gapop 10.0 , Gapext 0.5

Searched:     283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:   283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
              Maximum Match 100%
              Listing first 45 summaries

Database :     PIR_79:*
           1: pir1:*
           2: pir2:*
           3: pir3:*
           4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match Length | DB | ID | Description |
|---|---|---|---|---|---|
| 1 | 93 | 100.0 | 88 | 2 | PC4238 | heat shock protein |
| 2 | 93 | 100.0 | 94 | 2 | JC1479 | heat shock protein |
| 3 | 93 | 100.0 | 94 | 2 | A49855 | heat shock protein |
| 4 | 93 | 100.0 | 94 | 2 | A83720 | class I heat-shock |
| 5 | 91.4 | 94 | 2 | A41884 | heat shock protein |
| 6 | 90.3 | 94 | 2 | AE1333 | class I heat-shock |
| 7 | 90.3 | 94 | 2 | AD1704 | class I heat-shock |
| 8 | 84.9 | 122 | 2 | B70489 | GroES - Aquifex ae |
| 9 | 83.9 | 103 | 1 | BVYCGS | chaperonin groES - |
| 10 | 83.9 | 103 | 2 | A36721 | groES protein - Sy |
| 11 | 83.9 | 104 | 2 | S35310 | heat shock protein |
| 12 | 82.8 | 98 | 2 | JN0510 | heat shock protein |
| 13 | 82.8 | 101 | 2 | JN0513 | class I heat-shock |
| 14 | 81.7 | 95 | 2 | B71668 | 10 kd chaperonin ( |
| 15 | 81.7 | 98 | 1 | A43827 | chaperonin groES - |
| 16 | 81.7 | 98 | 2 | D95311 | groES2 chaperonin |
| 17 | 81.7 | 98 | 2 | AF3640 | 10K chaperonin gro |
| 18 | 80.6 | 103 | 2 | T06830 | chaperonin groES - |
| 19 | 79.6 | 104 | 2 | S35308 | heat shock protein |
| 20 | 78.5 | 103 | 2 | AF2263 | chaperonin GroES [ |
| 21 | 77.4 | 88 | 2 | A71253 | probable chaperoni |
| 22 | 77.4 | 88 | 2 | BVMY7B | probable chaperonin |
| 23 | 77.4 | 100 | 1 | BVMYBA | chaperonin groES - |
| 24 | 77.4 | 102 | 2 | A41325 | heat shock protein |
| 25 | 77.4 | 102 | 2 | S37565 | chaperonin groES - |
| 26 | 76.3 | 94 | 2 | S68248 | chaperonin groES |
| 27 | 76.3 | 94 | 2 | JC2563 | chaperonin groES - |
| 28 | 76.3 | 95 | 2 | C97442 | heat shock protein |
| 29 | 76.3 | 98 | 2 | AE2660 | 10 KD chaperonin ( |
| 30 | 76.3 | 98 | 2 | A36917 | heat shock protein |
| 31 | 76.3 | 99 | 2 | A97821 | 10K chaperonin [in |
| 32 | 76.3 | 71 | 2 | S77563 | chaperonin groES - |
| 33 | 76.3 | 106 | 2 | C71986 | 10Kda chaperone |
| 34 | 76.3 | 118 | 2 | S61396 | chaperonin groES - |
| 35 | 75.3 | 118 | 2 | G81019 | chaperonin 10 kDa |
| 36 | 75.3 | 96 | 2 | G81019 | chaperonin 10 Kd 8 |
| 37 | 75.3 | 96 | 2 | C87334 | chaperonin 10 kDa |
| 38 | 74.2 | 99 | 2 | S70668 | heat shock protein |
| 39 | 73.1 | 118 | 2 | I40330 | co-chaperonin GroE |
| 40 | 72.0 | 95 | 2 | A41872 | chaperonin - Deino |
| 41 | 72.0 | 95 | 2 | G97232 | chaperonin - Th |
| 42 | 72.0 | 95 | 2 | F75499 | groES protein (Gro |
| 43 | 71.0 | 92 | 2 | G72367 | chaperonin - Th |
| 44 | 67.7 | 86 | 2 | F81328 | 10 kD chaperonin ( |
| 45 | 67.7 | 253 | 2 | T52613 | chaperonin 21 prec |

ALIGNMENTS

RESULT 1
PC4238
heat shock protein GroES - Bacillus sp. (fragment)
N;Alternate names: molecular chaperone 60 GroES
C;Species: Bacillus sp.
C;Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C;Accession: PC4238; PC6023
R;Xu, Y.; Kobayashi, T.; Kudo, T.
Biosci. Biotechnol. Biochem. 60, 1633-1636, 1996
A;Residues: 1-88 <XUA>
A;Cross-references: UNIPROT:O50304; DDBJ:D55630; NID:g1682949; PIDN:BAA09493.1; PID:g1
A;Reference number: JC6063
A;Accession: PC6023
A;Molecule type: DNA
A;Residues: 1-88 <XU2>
A;Experimental source: C-125 strain BD224
C;Comment: This protein plays a role in protein folding by binding to the complex of a
C;Genetics:
A;Gene: groES
C;Superfamily: chaperonin groES
C;Keywords: molecular chaperone

| Qy | 1 ETEEKTASGIVLPDTAKEK 19 |
|---|---|
| | ||||||||||||||||||| |
| Db | 9 ETEEKTASGIVLPDTAKEK 27 |

Query Match     100.0%;  Score 93;  DB 2;  Length 88;
Best Local Similarity  100.0%;  Pred. No. 2.5e-08;
Matches  19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

RESULT 2
JC1479
heat shock protein TGroES - thermophilic bacterium PS-3
N;Alternate names: heat shock 12K protein; hsp10
C;Species: thermophilic bacterium PS-3
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 12-Sep-1997
C;Accession: JC1479; JC1194
R;Ohta, T.; Honda, K.; Saito, K.; Hayashi, H.; Tano, H.; Hamamoto, T.; Kagawa, Y.
Biochem. Biophys. Res. Commun. 191, 550-557, 1993
A;Title: Heat shock promoter of thermophilic chaperonin operon.
A;Reference number: JC1479; MUID:93213292; PMID:8096382
A;Accession: JC1479

| Qy | 1 ETEEKTASGIVLPDTAKEK 19 |
|---|---|
| | ||||||||||||||||||| |
| Db | 9 ETEEKTASGIVLPDTAKEK 27 |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:           September 29, 2005, 19:33:34 ; Search time 113.5 Seconds
                  (without alignments)
                  85.723 Million cell updates/sec

Title:            US-10-714-212-22
Perfect score:    93
Sequence:         1 ETEEKTASGIVLPDTAKEK 19

Scoring table:    BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:         1612278 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:   1612278

Minimum DB seq length:  0
Maximum DB seq length:  2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :        UniProt_03:*
                  1: uniprot_sprot:*
                  2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Accession | Description |
|---|---|---|---|---|---|---|---|
| 1 | 93 | 100.0 | 94 | 1 | CH10_BACHD | O50304 | bacillus ha |
| 2 | 93 | 100.0 | 94 | 1 | CH10_BACP3 | P26210 | bacillus ps |
| 3 | 93 | 100.0 | 94 | 1 | CH10_BACST | Q07200 | bacillus st |
| 4 | 93 | 100.0 | 94 | 1 | CH10_BACTR | Q8vv85 | bacillus th |
| 5 | 93 | 100.0 | 94 | 2 | CH10_BACTR | Q7dji2 | bacillus sp |
| 6 | 93 | 100.0 | 94 | 2 | CH10_BACSU | Q9ezv5 | bacillus st |
| 7 | 85 | 91.4 | 94 | 1 | CH10_BACSU | P28599 | bacillus su |
| 8 | 85 | 91.4 | 94 | 2 | Q63GV8 | Q63gv8 | bacillus ce |
| 9 | 85 | 91.4 | 94 | 2 | Q65M29 | Q65m29 | bacillus li |
| 10 | 85 | 91.4 | 94 | 2 | Q73ES0 | Q73es0 | bacillus ce |
| 11 | 85 | 91.4 | 94 | 2 | Q81VE2 | Q81ve2 | bacillus an |
| 12 | 85 | 91.4 | 94 | 2 | Q6HPC8 | Q6hpc8 | bacillus th |
| 13 | 85 | 91.4 | 96 | 2 | Q814B1 | Q814b1 | bacillus ce |
| 14 | 84 | 90.3 | 94 | 1 | CH10_LISIN | Q7ttx2 | listeria in |
| 15 | 84 | 90.3 | 94 | 1 | CH10_LISMO | Q71xu5 | listeria mo |
| 16 | 84 | 90.3 | 94 | 1 | CH10_LISMO | Q8age7 | listeria mo |
| 17 | 84 | 90.3 | 94 | 2 | Q929u9 | Q929u9 | listeria in |
| 18 | 82 | 88.2 | 103 | 1 | CH10_SYNP7 | Q7tus3 | synechococc |
| 19 | 82 | 88.2 | 103 | 2 | Q7tv92 | Q7tv92 | prochloroco |
| 20 | 82 | 88.2 | 103 | 2 | Q98il0 | Q98il0 | prochloroco |
| 21 | 81 | 87.1 | 104 | 1 | CH12_RHILO | Q6w164 | rhizobium l |
| 22 | 81 | 87.1 | 104 | 2 | Q6wid4 | Q6wid4 | rhizobium s |
| 23 | 81 | 87.1 | 104 | 2 | Q67kb7 | Q67kb7 | symbiobacte |
| 24 | 79 | 84.9 | 97 | 2 | CH10_AQUAE | O67942 | aquifex aeo |
| 25 | 79 | 84.9 | 122 | 1 | CH10_SYNP6 | P07889 | synechococc |
| 26 | 78 | 83.9 | 103 | 1 | CH10_SYNP7 | P22880 | synechococc |
| 27 | 78 | 83.9 | 103 | 2 | Q7tu43 | Q7u43 | prochloroco |
| 28 | 78 | 83.9 | 104 | 1 | P77828 | P77828 | bradyrhizob |
| 29 | 78 | 83.9 | 104 | 1 | CH12_RHOPA | P60367 | rhodopseudo |
| 30 | 78 | 83.9 | 104 | 2 | CH13_BRAJA | P35864 | bradyrhizob |
| 31 | 78 | 83.9 | 104 | 2 | CH11_BRAJA | Q89lk9 | bradyrhizob |
| 32 | 78 | 83.9 | 116 | 2 | Q8gb96 | Q8gb96 | heliobacill |
| 33 | 78 | 83.9 | 93 | 1 | CH10_OCEIH | Q8cx14 | oceanobacil |
| 34 | 77 | 82.8 | 98 | 1 | CH11_RHIME | P35473 | rhizobium m |
| 35 | 77 | 82.8 | 98 | 1 | CH13_RHILO | Q98ax8 | rhizobium l |
| 36 | 77 | 82.8 | 98 | 1 | CH14_RHILO | Q983s3 | rhizobium l |
| 37 | 77 | 82.8 | 98 | 1 | CH15_RHILO | Q98lk0 | rhizobium l |
| 38 | 77 | 82.8 | 98 | 2 | Q799q4 | Q799q4 | rhizobium l |
| 39 | 77 | 82.8 | 98 | 2 | Q6fyu9 | Q6fyu9 | rhizobium l |
| 40 | 77 | 82.8 | 98 | 2 | Q6fyu9 | Q6fyu9 | bartonella |
| 41 | 77 | 82.8 | 103 | 1 | Q6H8L3 | Q6hbl3 | bartonella |
| 42 | 77 | 82.8 | 104 | 1 | CH10_SYNP8 | P35474 | synechococc |
| 43 | 76 | 81.7 | 94 | 1 | O81373 | Q81373 | synechococc |
| 44 | 76 | 81.7 | 95 | 1 | CH10_RICTY | Q9ct6 | rickettsia |
| 45 | 76 | 81.7 | 95 | 2 | CH10_RICPR | Q72al5 | desulfovibr |

ALIGNMENTS

RESULT 1
CH10_BACHD
ID   CH10_BACHD    STANDARD;    PRT;    94 AA.
AC   O50304;
DT   30-MAY-2000 (Rel. 39, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   10 kDa chaperonin (Protein Cpn10) (groES protein)
GN   Name=groS; Synonyms=groES, mopB; OrderedLocusNames=BH0561;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE OF 7-94 FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=97141316; PubMed=8987660;
RA   Xu Y., Kobayashi T., Kudo T.;
RT   "Molecular cloning and nucleotide sequence of the groEL gene from the
RT   alkaliphilic Bacillus sp. strain C-125 and reactivation of thermally
RT   inactivated alpha-glucosidase by recombinant GroEL.";
RL   Biosci. Biotechnol. Biochem. 60:1633-1636(1996).
CC   -!- FUNCTION: Binds to Cpn60 in the presence of Mg-ATP and suppresses
CC       the ATPase activity of the latter.
CC   -!- SUBUNIT: Heptamer of 7 subunits arranged in a ring (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- INDUCTION: By heat shock.
CC   -!- SIMILARITY: Belongs to the groES chaperonin family.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; AP001508; BAB04280.1; -.
DR   EMBL; D55630; BAA09493.1; -.
DR   PIR; A83720; A83720.
DR   PIR; PC4238; PC4238.
DR   HSSP; P09621; 1HX5.
DR   HAMAP; MF_00580; -; 1.
DR   InterPro; IPR001476; Chaprnin_Cpn10.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: September 29, 2005, 19:35:05 ; Search time 30.5 Seconds
(without alignments)
46.503 Million cell updates/sec

Title:         US-10-714-212-22
Perfect score: 93
Sequence:      1 ETEEKTASGIVLPDTAKEK 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:      513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:          513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 93 | 100.0 | 19 | 4 | US-09-470-830A-20 | Sequence 20, Appl |
| 2 | 93 | 100.0 | 19 | 4 | US-09-470-830A-22 | Sequence 22, Appl |
| 3 | 88 | 94.6 | 19 | 2 | US-08-467-822-38 | Sequence 38, Appl |
| 4 | 88 | 94.6 | 94 | 3 | US-08-432-697-38 | Sequence 38, Appl |
| 5 | 85 | 91.4 | 94 | 3 | US-08-466-248-38 | Sequence 38, Appl |
| 6 | 85 | 91.4 | 19 | 4 | US-09-470-830A-21 | Sequence 21, Appl |
| 7 | 78 | 83.9 | 19 | 4 | US-09-470-830A-24 | Sequence 24, Appl |
| 8 | 78 | 83.9 | 19 | 4 | US-09-470-830A-29 | Sequence 29, Appl |
| 9 | 78 | 83.9 | 19 | 4 | US-09-470-830A-29 | Sequence 29, Appl |
| 10 | 77 | 82.8 | 19 | 4 | US-09-470-830A-30 | Sequence 30, Appl |
| 11 | 77 | 82.8 | 19 | 4 | US-09-470-830A-10 | Sequence 10, Appl |
| 12 | 76 | 81.7 | 19 | 4 | US-09-470-830A-17 | Sequence 17, Appl |
| 13 | 74 | 79.6 | 19 | 4 | US-09-470-830A-13 | Sequence 13, Appl |
| 14 | 74 | 79.6 | 19 | 4 | US-09-470-830A-19 | Sequence 19, Appl |
| 15 | 72 | 77.4 | 19 | 4 | US-09-470-830A-23 | Sequence 23, Appl |
| 16 | 72 | 77.4 | 19 | 4 | US-09-470-830A-26 | Sequence 26, Appl |
| 17 | 72 | 77.4 | 198 | 4 | US-09-613-303-35 | Sequence 35, Appl |
| 18 | 72 | 77.4 | 198 | 4 | US-10-267-311-35 | Sequence 35, Appl |
| 19 | 71 | 76.3 | 19 | 4 | US-09-470-830A-14 | Sequence 14, Appl |
| 20 | 71 | 76.3 | 198 | 2 | US-08-467-822-29 | Sequence 29, Appl |
| 21 | 71 | 76.3 | 118 | 2 | US-08-432-697-29 | Sequence 29, Appl |
| 22 | 71 | 76.3 | 118 | 3 | US-08-466-248-29 | Sequence 29, Appl |
| 23 | 71 | 76.3 | 118 | 3 | US-09-470-830A-31 | Sequence 31, Appl |
| 24 | 68 | 73.1 | 99 | 4 | US-09-603-208A-4 | Sequence 4, Appl |
| 25 | 67 | 72.0 | 19 | 4 | US-09-470-830A-15 | Sequence 15, Appl |
| 26 | 61.5 | 66.1 | 102 | 4 | US-09-438-185A-137 | Sequence 137, Appl |
| 27 | 61 | 65.5 | 19 | 4 | US-09-470-830A-25 | Sequence 25, Appl |
| 28 | 61 | 65.6 | 111 | 3 | US-09-134-001C-4686 | Sequence 4686, Ap |
| 29 | 60.5 | 65.1 | 102 | 4 | US-09-198-4452A-154 | Sequence 154, Appl |
| 30 | 60 | 64.5 | 19 | 4 | US-09-470-830A-6 | Sequence 6, Appli |
| 31 | 60 | 64.5 | 19 | 4 | US-09-470-830A-9 | Sequence 9, Appli |
| 32 | 60 | 64.5 | 19 | 4 | US-09-470-830A-11 | Sequence 11, Appl |
| 33 | 60 | 64.5 | 94 | 2 | US-08-467-822-37 | Sequence 37, Appl |
| 34 | 60 | 64.5 | 94 | 2 | US-08-432-697-37 | Sequence 37, Appl |
| 35 | 60 | 64.5 | 94 | 3 | US-08-466-248-37 | Sequence 37, Appl |
| 36 | 60 | 64.5 | 96 | 3 | US-09-470-830A-27 | Sequence 27, Appl |
| 37 | 60 | 64.5 | 140 | 4 | US-09-540-236-2611 | Sequence 2611, Ap |
| 38 | 58 | 62.4 | 145 | 4 | US-09-225-9901A-24284 | Sequence 24284, A |
| 39 | 57.5 | 61.8 | 19 | 4 | US-09-470-830A-32 | Sequence 32, Appl |
| 40 | 57 | 61.3 | 19 | 4 | US-09-470-830A-27 | Sequence 27, Appl |
| 41 | 57 | 61.3 | 93 | 2 | US-08-467-822-36 | Sequence 36, Appl |
| 42 | 57 | 61.3 | 93 | 2 | US-08-432-697-36 | Sequence 36, Appl |
| 43 | 57 | 61.3 | 93 | 3 | US-08-466-248-36 | Sequence 36, Appl |
| 44 | 57 | 61.3 | 106 | 4 | US-09-328-352-5510 | Sequence 5510, Ap |
| 45 | 56 | 60.2 | 19 | 4 | US-09-470-830A-8 | Sequence 8, Appl |

ALIGNMENTS

RESULT 1
US-09-470-830A-20
; Sequence 20, Application US/09470830A
; Patent No. 6677139
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; TITLE OF INVENTION: Host Cells
; FILE REFERENCE: GC559-PCT
; CURRENT APPLICATION NUMBER: US/09/470,830A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/470,830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Thermoactinomyces sp.
US-09-470-830A-20

Query Match 100.0%; Score 93; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1  ETEEKTASGIVLPDTAKEK 19
        |||||||||||||||||||
Db  1  ETEEKTASGIVLPDTAKEK 19

RESULT 2
US-09-470-830A-22
; Sequence 22, Application US/09470830A
; Patent No. 6677139
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; TITLE OF INVENTION: Host Cells
; FILE REFERENCE: GC559-PCT
; CURRENT APPLICATION NUMBER: US/09/470,830A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/470,830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-09-470-830A-22

OM protein - protein search, using sw model

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Run on: September 29, 2005, 19:38:05 ; Search time 117 Seconds
(without alignments)
67.412 Million cell updates/sec

Title:          US-10-714-212-22
Perfect score:  93
Sequence:       1 ETEEKTASGIVLPDTAKEK 19
Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched:       1846076 seqs, 41516000 residues

Total number of hits satisfying chosen parameters:    1846076

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%;
                 Maximum Match 100%;
                 Listing first 45 summaries

Database :  Published Applications AA:*

1:  /cgn2_6/ptcdata/1/pubpaa/US07_PUBCOMB.pep:*
2:  /cgn2_6/ptcdata/1/pubpaa/PCT_NEW_PUB.pep:*
3:  /cgn2_6/ptcdata/1/pubpaa/US09C_PUBCOMB.pep:*
4:  /cgn2_6/ptcdata/1/pubpaa/US09_NEW_PUB.pep:*
5:  /cgn2_6/ptcdata/1/pubpaa/US06_PUBCOMB.pep:*
6:  /cgn2_6/ptcdata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7:  /cgn2_6/ptcdata/1/pubpaa/US07_NEW_PUB.pep:*
8:  /cgn2_6/ptcdata/1/pubpaa/US08_PUBCOMB.pep:*
9:  /cgn2_6/ptcdata/1/pubpaa/US08_NEW_PUB.pep:*
10: /cgn2_6/ptcdata/1/pubpaa/US09A_PUBCOMB.pep:*
11: /cgn2_6/ptcdata/1/pubpaa/US09B_PUBCOMB.pep:*
12: /cgn2_6/ptcdata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptcdata/1/pubpaa/US0A_PUBCOMB.pep:*
14: /cgn2_6/ptcdata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptcdata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptcdata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptcdata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptcdata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptcdata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptcdata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptcdata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptcdata/1/pubpaa/US60_PUBCOMB.pep:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 93 | 100.0 | 19 | 15 | US-10-714-212-20 | Sequence 20, Appl |
| 2 | 93 | 100.0 | 19 | 15 | US-10-714-212-22 | Sequence 22, Appl |
| 3 | 85 | 91.4 | 19 | 15 | US-10-369-493-21 | Sequence 21, Appl |
| 4 | 81 | 87.1 | 95 | 15 | US-10-369-493-17908 | Sequence 17908, A |
| 5 | 81 | 87.1 | 104 | 15 | US-10-369-493-11921 | Sequence 11921, A |
| 6 | 79 | 84.9 | 122 | 15 | US-10-369-493-166 | Sequence 166, App |
| 7 | 78 | 83.9 | 19 | 15 | US-10-714-212-24 | Sequence 24, Appl |
| 8 | 78 | 83.9 | 19 | 15 | US-10-714-212-29 | Sequence 29, Appl |
| 9 | 78 | 83.9 | 19 | 15 | US-10-714-212-30 | Sequence 30, Appl |
| 10 | 78 | 83.9 | 104 | 15 | US-10-369-493-20711 | Sequence 20711, A |
| 11 | 77 | 82.8 | 19 | 15 | US-10-369-493-10 | Sequence 10, Appl |
| 12 | 77 | 82.8 | 19 | 15 | US-10-714-212-17 | Sequence 17, Appl |
| 13 | 77 | 82.8 | 94 | 15 | US-10-369-493-10863 | Sequence 10863, A |
| 14 | 77 | 82.8 | 98 | 15 | US-10-369-493-12331 | Sequence 12331, A |
| 15 | 76 | 81.7 | 19 | 15 | US-10-714-212-12 | Sequence 12, Appl |
| 16 | 75 | 80.6 | 89 | 15 | US-10-369-493-10294 | Sequence 10294, A |
| 17 | 74 | 79.6 | 96 | 15 | US-10-714-212-13 | Sequence 13, Appl |
| 18 | 73 | 78.5 | 19 | 15 | US-10-369-493-10015 | Sequence 10015, A |
| 19 | 72 | 77.4 | 19 | 15 | US-10-714-212-19 | Sequence 19, Appl |
| 20 | 72 | 77.4 | 19 | 15 | US-10-714-212-23 | Sequence 23, Appl |
| 21 | 72 | 77.4 | 19 | 15 | US-10-714-212-26 | Sequence 26, Appl |
| 22 | 72 | 77.4 | 95 | 15 | US-10-369-493-21767 | Sequence 21767, A |
| 23 | 72 | 77.4 | 100 | 9 | US-09-712-363-272 | Sequence 272, App |
| 24 | 72 | 77.4 | 100 | 16 | US-10-332-512A-11 | Sequence 11, Appl |
| 25 | 72 | 77.4 | 102 | 14 | US-10-156-761-12525 | Sequence 12525, A |
| 26 | 72 | 77.4 | 198 | 14 | US-10-267-311-35 | Sequence 35, Appl |
| 27 | 72 | 77.4 | 198 | 17 | US-10-679-956-35 | Sequence 35, Appl |
| 28 | 71 | 76.3 | 19 | 15 | US-10-714-212-14 | Sequence 14, Appl |
| 29 | 71 | 76.3 | 19 | 17 | US-10-714-212-31 | Sequence 31, Appl |
| 30 | 71 | 76.3 | 71 | 15 | US-10-369-493-11759 | Sequence 11759, A |
| 31 | 71 | 76.3 | 98 | 15 | US-10-369-493-14206 | Sequence 14206, A |
| 32 | 71 | 76.3 | 98 | 15 | US-10-369-493-14459 | Sequence 14459, A |
| 33 | 71 | 76.3 | 98 | 15 | US-10-369-493-14919 | Sequence 14919, A |
| 34 | 71 | 76.3 | 98 | 15 | US-10-369-493-15264 | Sequence 15264, A |
| 35 | 71 | 76.3 | 118 | 9 | US-09-815-242-11281 | Sequence 11281, A |
| 36 | 71 | 76.3 | 118 | 9 | US-09-815-242-11465 | Sequence 11465, A |
| 37 | 71 | 76.3 | 118 | 18 | US-10-953-901-419 | Sequence 419, App |
| 38 | 71 | 76.3 | 118 | 18 | US-10-953-901-421 | Sequence 421, App |
| 39 | 69 | 74.2 | 95 | 15 | US-10-369-493-7916 | Sequence 7916, Ap |
| 40 | 69 | 74.2 | 96 | 15 | US-10-369-493-16818 | Sequence 16818, A |
| 41 | 68 | 73.1 | 99 | 15 | US-09-738-626-4164 | Sequence 4164, Ap |
| 42 | 67 | 72.0 | 19 | 15 | US-10-714-212-15 | Sequence 15, Appl |
| 43 | 67 | 72.0 | 94 | 15 | US-10-369-493-15645 | Sequence 15645, A |
| 44 | 67 | 72.0 | 94 | 15 | US-10-369-493-16020 | Sequence 16020, A |
| 45 | 67 | 72.0 | 94 | 15 | US-10-369-493-16372 | Sequence 16372, A |

ALIGNMENTS

RESULT 1
US-10-714-212-20
; Sequence 20, Application US/10714212
; Publication No. US20040077038A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; TITLE OF INVENTION: Host Cells
; FILE REFERENCE: GC559-PCT
; CURRENT APPLICATION NUMBER: US/10/714,212
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 09/470,830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Thermoactinomyces sp.
US-10-714-212-20

Query Match 100.0%;  Score 93;  DB 15;  Length 19;
Best Local Similarity 100.0%;
Matches 19;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy 1 ETEEKTASGIVLPDTAKEK 19
      |||||||||||||||||||
Db 1 ETEEKTASGIVLPDTAKEK 19

RESULT 2
US-10-714-212-22
; Sequence 22, Application US/10714212
; Publication No. US20040077038A1